In the Specification:

Please amend the specification as shown:

Please delete the paragraph on page 2, line 25 to page 5, line 14 and replace it with the following paragraph:

This aspect of the present invention provides for libraries of proteins, especially recombinant antibody domains such as Fv's, whereby individual protein members of the library include, within their amino acid sequence, a tract of sequence (a "barcode") which can subsequently be sequenced in order to identify which protein(s) has bound to the specific target (or, in the case of Fv's, "antigen"). This embodiment will apply especially where the Fv's are derived from human genes whereby the selected Fv may be suitable for human therapeutic or diagnostic use. In this particular application, an extensive gene library of Fv's is created from a pool of immunoglobulin cDNA's such as those derived from peripheral blood B cells in humans or such as pools created synthetically using human variable regions with semi-randomised ("combinatorial") CDRs (complimentarity-determining regions) at one or more positions. If this gene library is created in such manner that a random (or semi-random) gene sequence is included within the Fv coding region or terminal to this region, then such a random/semi-random gene sequence will generate a random/semi-random peptide sequence associated with individual Fv's. Such a random/semi-random gene sequence is created using standard methods such as oligonucleotide priming/DNA polymerase extension or PCR whereby a random/semi-random synthetic oligonucleotide sequence is used as one of a pair of primers used to amplify immunoglobulin gene fragments during the creation of the Fv gene library. If members of the Fv library comprise two chains (i. e. heavy and light chain-derived chains (VH and VL)) as opposed to a single-chain (VH and VL joined by a peptide linker), then individual barcodes can be associated with each of the chains (or can be associated with one of the chains only). Upon creation of the library, the resultant Fv's each include one or more "peptide barcodes" unique to that particular Fv or to a small subset of Fv's from within the complex library. Preferably, the peptide barcode is C terminal to the single-chain Fv region or C terminal to the VH or VL or both and includes, flanked between itself and the Fv region, one or more protease sensitive sites such as sites for enterokinase (cleaves after Asp-Asp-Asp-Asp-Lys **(SEQ ID NO: 1)**, Factor Xa (cleaves after Ile-Glu/Asp-Gly-Arg) or other endopeptidases. If a mixture of such Fv's is produced from a

suitable gene library, then this mixture is mixed with a target antigen (or antigens such as on cells), usually where the antigen is immobilised. This results in specific Fv's binding to the target antigen with non-binders (or weak binders depending on the stringency of washing) being washed away. Having washed away excess antibodies, the remaining antigen/Fv complex is then usually released from the Fv by digestion with the endoprotease used to cleave the introduced protease sensitive site. This released barcoded peptide is then subjected to mass analysis / mass spectrometry sequencing either directly or, if desired, following capture by virtue of specific amino acids or amino acid sequences which allow the peptide to be captured onto a solid phase such a cysteine residues which can be biotinylated for subsequent capture on immobilised avidin or streptavidin. Alternatively, any other method can be employed to determine the sequence of the peptide barcode either within the Fv or after release including using specific ligands which bind to the barcode in a sequence-specific manner. Having determined the sequences (or part-sequence) of barcodes derived from bound Fv's, corresponding synthetic oligonucleotides are then produced and used to specifically amplify or enrich for specific Fv genes from the library. These specific or enriched Fv (or VH and VL) genes are then further used to generate corresponding Fv's which could then be retested for antigen binding either individually or as part of a small pool of isolated Fv's. Ultimately, by this method, specific Fv's can be generated with desirable antigen binding properties and, if from a human source, potential clinical utility. This aspect also encompasses the use of multiple barcodes associated with individual proteins or Fv's, for example two adjacent barcodes at the C terminus of Fv's whereby two peptides are released from each Fv by protease digestion, either simultaneously in order to enhance the identity of Fv's which bind to the target, or sequentially whereby different proteases are used in successive rounds of digestion to provide a different means to subsequently amplify Fv genes corresponding to Fv's which bind to the target. This aspect also encompasses the use of multiple barcodes which are analysed at the same time in order to increase the diversity of overall barcode sequences to provide specific coding of individual proteins. This aspect also encompasses the use of barcodes within individual proteins, for example within one or more CDR positions of an Fv. This aspect also encompasses the use of proteases which might also digest the protein components of the protein:target mixture or, additionally, any protein agent used to immobilise the target, with the proviso that the barcode peptides released from the bound test protein(s) can still be detected and sequenced within the background of other peptides. In the preferred format of this aspect, a single region of barcode is provided at the C

terminus of the light chains forming a soluble Fab fragment whereby VHs and VLs are encoded by the same expression cassette or cistron such that the barcode sequence can be used to access both VH and VL genes. Such an Fab fragment can be conveniently produced using a range of expression systems, for example the M13 bacteriophage vector system where, by introduction of secretory leader sequences, the heavy and light chains of Fabs are secreted into the periplasmic space of the host bacteria and harvested from that space. The vector system is first prepared with in-frame barcodes by cloning in mixtures of synthetic oligonucleotides. For the formation of two adjacent barcodes, this is conveniently undertaken by sequential cloning or oligonucelotide mutagenesis whereby pooled M13 recombinants containing the first mixed barcode are prepared as a template for subsequent cloning of the second in-frame barcode. Preferably, the barcoding is designed such that the encoded protein contains endonuclease sites both flanking and between the two barcodes and also whereby a "spacer" region adjacent to one of the barcodes creates a peptide including that barcode which has a higher molecular weight than the other barcode. By judicious design of barcodes and the use of multiple barcodes in this manner, there is provided an option to simply analyse masses of endoprotease-released peptides by, for example, MALDI-ToF whereby the sequences of the peptides can be deduced (or near deduced) such that synthetic oligonucleotides can be designed to isolate (or enrich) for the specific proteins with the barcode(s) detected by MALDI-ToF analysis. Such deduction of these sequences is achieved by design of sequences whereby specific amino acids only occur in one or two positions along the peptide. For example, where the peptide is designed using 17 of the 20 natural amino acids (hereby designated A-Q), then the sequences might be designed with options for any of three amino acids at each position along the peptide sequence as follows;

aa position:	1	2	3	4	5	6	7	8
amino acid options:	A	C	E	G	I	K	M	O
	B	D	F	H	J	L	N	P
	C	E	G	I	K	M	O	Q

This design would give a theoretical 6561 different peptide sequence barcodes. If an adjacent barcode with a spacer region is also designed on the same basis, then this would give an additional 6561 different barcodes. In combination, this would create 4×10^7 barcode sequences which would be adequate to uniquely tag most members of a protein library of such size. The use of additional adjacent barcodes or longer barcodes based, for example, on use of two specific amino acids at any position in the sequence (thus creating

262,144 different barcode sequences using 19 amino acids) would increase the diversity of barcodes provided. In practice, codon redundancy is reduced through the judicious choice of codons at each position in the sequence during design of mixed synthetic oligonucleotides. One design of oligonucleotide for an 8 amino acid barcode peptide for MS/MS sequencing is as follows;

Codons-	NAC	NCC	NGG	NTG	TKC	VAG	GNV	CNT	**(SEQ ID NO: 2)**
Amino acids-	N	T	R	L	F	Q	D	H	**(SEQ ID NO: 3)**
	D	P	G	M	C	E	V	L	
	H	A	W	V		K	A	P	
	Y	S					G	R	
							E		

where codons N = A, C, G or T
K = G or T
V = A, C or G

4 X 4 X 3 X 3 X 2 X 3 X 4 X 4 = 13824 barcode sequences

Please delete the paragraph on page 23, line 1 to page 24, line 15 and replace it with the following paragraph:

The fifth aspect of the present invention includes the use of protein affinity reagents other than monoclonal antibodies where such reagents can facilitate the fractionation of peptides or proteins prior to mass analysis. Such affinity reagents would include molecules of the immune which selectively bind certain peptides such as major histocompatability proteins and T cell receptors. Other protein affinity reagents would include protein domains commonly involved in protein-protein binding interactions such as SH1 domains. Included in the present invention is the concept of cyclising peptides including within mixtures and especially when bound to solid phases by, for example, linking cysteine residues under reducing conditions. One method for this would be to add an additional cysteine residue at an exposed N or C terminal on immobilised peptides using, for example for C terminal immobilised peptides, standard conditions of peptide synthesis or using reverse proteolysis whereby certain proteases such as carboxypeptidase Y and lysyl endopeptidase. Included in the fifth aspect is also a method for further fractionating proteins or peptides by adding, usually at the N terminus, amino acids which form part of the recognition sequence of a protease which specifically cleaves at a recognition sequence of two or amino acids whereby one or more terminal amino acids in the protease recognition site is provided by the starting protein or peptide. In this manner, only a fraction of the proteins or peptides to which

the new amino acids are added will be then subject to terminal protease cleavage by virtue of the newly created sequence. In this manner, proteins or peptides can be tagged with additional amino acids usually at the N terminus creating, in a fraction of the thus tagged mixture, a specific protease cleavage site. The proteins or peptides can then, for example, be immobilised via the new terminus for example using a tagged terminal amino acid or by adding a chemical tag to the terminus, whereby an affinity reagent is then used to immobilise the tagged moieties. After removing non-immobilised untagged molecules, the proteins or peptides can then be subjected to cleavage with the specific protease which will then only cleave where the cleavage site has been generated by a combination of synthesis-derived amino acids and the original protein or peptide-derived amino acids. The cleaved peptides can then be fractionated using protein affinity reagents and mass analysed (or further processed prior to mass analysis) thus representing a subset of the peptide mixture. By using parallel synthesis of specific amino acids to exposed termini followed by immobilisation and cleavage, large mixtures of proteins or peptides can be fractionated on the basis of their terminal amino acid(s). An example of a protease recognition site is ile, glu, gly, arg **(SEQ ID NO: 4)** which is cleaved between gly and arg by Factor Xa. The sequence ile, glu, gly could be synthesised onto the N terminus of a protein or peptide and thus if the adjacent amino acid in the protein or peptide sequence were arg, the cleavage site would be created and could be cleaved by Factor Xa. Other examples of protease cleavage sites are asp, asp, asp, asp, lys **(SEQ ID NO: 1)**, cleaved by Enterokinase between asp and lys; pro, gly, ala, ala, his, tyr **(SEQ ID NO: 5)** cleaved between his and tyr by genease I; leu, val, pro, arg, gly, ser **(SEQ ID NO: 6)** cleaved between arg and gly by thrombin. N terminal addition of partial sequence asp, asp, asp, asp **(SEQ ID NO: 7)** could be used to identify proteins or peptides with N terminal lys (cleaved by enterokinase), pro, gly, ala, ala, his **(SEQ ID NO: 8)** to identify proteins/peptides with N terminal tyr (cleaved by genease), leu, val, pro, arg **(SEQ ID NO: 9)** to identify N terminal gly, ser; or leu, val, pro, arg, gly **(SEQ ID NO: 10)** to identify N terminal ser (cleaved by thrombin). Other proteases such as the MMP's (matrix metalloproteinases) with specific recognition sites could be used to fractionate proteins with other N terminal amino acids. Different protease recognition sites could thus be used in combination with the proteases to fractionate proteins or peptides according to the N terminal amino acid. As an alternative, one or more amino acids are added to the free N terminus of a peptide could be used to create a site for binding by an affinity reagent including where such a site is dependant on one or more the N terminal amino acids from the peptide. Thus, different peptide or groups of peptides could be

distinguished by the addition of amino acids to the N terminus which creates, in a manner dependant on the N terminal amino acids, a site for protease digestion or a site for binding by an affinity reagent. Where proteins are used as the starting material especially from mammalian cells whereby the N terminal protein is methionine, this can be removed if required by, for example, formylation and cleavage by a bacterial protease specific for removal of terminal formylmethionine.

Please delete the paragraphs on page 26, line 4 to page 28, line 16 and replace them with the following paragraphs:

The experiments described in the present example were conducted using a pair of modified single chain antibody (scAbs) genes. Two modified scAbs were prepared consisting of N-terminal epitope tags, the heavy chain variable region (VH), a 14 amino acid linker (EGKSSGSGSESKVD) **(SEQ ID NO: 11)**, the light chain variable region (VL) each fused to the b-zip domain from either the c-jun or c-fos genes.

These constructs were cloned into the vector pET 5c (Rosenberg AH et al., Gene, 56: 125-135, 1987) which provides a T7 promoter followed by the ribosome binding site from T7 gene 10. The scAb constructs were inserted into the vector at an NdeI site such that the sequence encoding the epitope tag followed the first ATG of T7 gene 10. The first construct consisted of a scAb against Pseudomonas aeruginosa (Molloy P. et al. *Journal of Applied Bacteriology*, 78: 359-365, 1995) with the FLAG epitope (MDYKDDDK) **(SEQ ID NO: 12)** (Knappik A and Pluckthun A, *BioTechniques*, **17**: 754-761, 1994) added at the N terminus, and the b-zip domain of c-fos (Abate, C. et al *Proc. Natl. Acad. Sci.* USA. **87:** 1032-1036, 1990) at the C-terminal region of the protein. The second consisted a scAb constructed from the anti-foetal antigen antibody 340 (Durrant LG et al. *Prenatal Diagnosis*, 14: 131-140, 1994) with a poly-Histidine tag at the N terminus, and the b-zip domain of c-jun (Abate C. et al, *ibid*) at the C-terminal region of the protein.

The anti-*Pseudomonas aeruginosa* (α-Ps-fos) scAb and the 340-jun scAb were constructed as described below:

DNA for the α-Ps scAb in the vector pPMIHis (Molloy P et al., *ibid*) was amplified with the primers RD 5'FLAG: 5'gcggatcccatatggactacaaagacgatgacgacaaacaggtgcagctgcag3' **(SEQ ID NO: 13)** (Genosys Biotechnologies Europe Ltd, Cambridge, UK) and RD 3' :

5'gcgaattcgtggtggtggtggtgtgactctcc3' **(SEQ ID NO: 14)** (Genosys) which introduced the 5'FLAG epitope sequence and removed the 3' stop codon respectively. The reaction mixture included 0.1 µg template DNA, 2.6 units of Expand™ High Fidelity PCR enzyme mix (Boehringer Mannheim, Lewes, UK.), Expand HF buffer (Boehringer Mannheim), 1.5 mM $MgCl_2$, 200 µM M deoxynucleotide triphosphates (dNTPs) (Life Technologies, Paisley, UK) and 25 pmoles of each primer. Cycles were 96°C 5 minutes, followed by [95°C 1 minute, 50°C 1 minute, 72°C 1 minute] times 5, [95°C 45 seconds, 50°C 1 minute, 72°C 1 minute 30 seconds] times 8, [95°C 45 seconds, 50°C 1 minute, 72°C 2 minutes] times 5, finishing with 72°C 5 minutes. The 1123 bp product obtained was cut with BamHI and EcoRI and cloned into the vector pUCI9 (Boehringer Mannheim). The DNA sequence was confirmed, using the Thermo Sequenase radiolabeled terminator cycle sequencing kit with [^{33}P] dideoxy nucleotides (Amersham Life Science, Amersham, UK). The construct was cloned into pET5c vector (Promega UK Ltd, Southampton, UK.) as a NdeI to EcoRI fragment (see *Molecular Cloning, A Laboratory Manual* eds. Sambrook J, fritsch EF, Maniatis T. Cold Spring Harbor Laboratory Press 1989, New York, USA). Plasmid DNA was prepared using Wizard® Plus SV Minipreps DNA purification System (Promega UK Ltd), or for larger scale, Qiagen Plasmid Midi Kit (Qiagen Ltd, Crawley, UK.). The new plasmid generated was named pET5c FLAG-αPs scAb.

The fos cassette was assembled by PCR of overlapping oligonucleotides:

Fos1for 5'-atggaattcctcgagaccgacaccctacaggcggaaaccgaccagctgga **(SEQ ID NO: 15)**

Fos80rev 5'-tcgcgatttcggtttgcagcgcggatttttcgtcttccagctggtcggtt **(SEQ ID NO: 16)**

Fos 7Ifor 5'-aaaccgaaatcgcgaacctgctgaaagaaaaagaaaagctggagttcatc **(SEQ ID NO: 17)**

Fos 155rev 5'ggaagcttgaattccgccggacggtgtgccgccaggatgaactccagctt **(SEQ ID NO: 18)**

The above oligonucleotides were included in a reaction mix at 1pmol each, and the reaction was driven using 10pmol primers Fos1fS; 5'-atggaattcctcgagacc **(SEQ ID NO: 19)** and Fos 155rS 5'-ggaagcttgaattccgcc **(SEQ ID NO: 20)** using high fidelity polymerase and reaction components as previously. The resulting 155bp product was digested with EcoRI, purified and cloned into EcoRI cut pUC19 for sequence analysis using standard procedures (see *Molecular Cloning, A Laboratory Manual* ibid). The Fos cassette was sub-cloned into the

pET5c FLAG-αPs scAb plasmid as an XhoI-EcoRI fragment by substitution of the existing 320bp XhoI-EcoRI fragment carrying the human constant region domain.

The 340 scAb was produced by substitution the VH and VK of the 340 antibody in place of the α-Ps VH and VK in ppM1His. The 340 VH was amplified with the primers 5'cagctgcaggagtctgggggaggcttag3' **(SEQ ID NO: 21)** (Genosys) and 5'tcagtagacggtgaccgaggttccttgaccccagta3' **(SEQ ID NO: 22)** (Genosys). The reaction mixture included 0.1 µg template DNA, 2.6 units of Expand™ High Fidelity PCR enzyme mix, Expand HF buffer, 1.5 mM MgCl2,200 µM dNTPs and 25 pmoles of each primer. Cycles were 96°C 5 minutes, followed by [95°C 1 minute, 50°C 1 minute, 72°C 1 minute] times 5, [95°C 45 seconds, 50°C 1 minute, 72°C 1 minute 30 seconds] times 8, [95°C 45 seconds, 50°C 1 minute, 72°C 2 minutes] times 5, finishing with 72°C 5 minutes. The 357 bp product was cut with PstI and BstEII and cloned into PstI and BstEII cut pPM1His (see *Molecular Cloning, A Laboratory Manual*, ibid). Similarly, the 340 VK was amplified with the primers 5'gtgacattgagctcacacagtctcct3' **(SEQ ID NO: 23)** and 5'cagcccgttttatctcgagcttggtccg3' **(SEQ ID NO: 24)** (Genosys). The 339 bp product was cut with SstI and XhoI and cloned into SstI and XhoI cut modified pPM1His (produced above). The DNA sequence was confirmed, using the Thermo Sequenase radiolabeled terminator cycle sequencing kit with [^{33}P] dideoxy nucleotides as before. DNA for the 340 scAb in the vector pPM1His was amplified with the primers RD 5' HIS: 5'gcggatcccatatgcaccatcatcaccatcaccaggtgcagctgcag3' **(SEQ ID NO: 25)** (Genosys) and RD 3' (given above) which introduced the 6 histidine residues at the 5' end and removed the 3' stop codon respectively. Reagents and conditions for amplification were exactly as for the α-Ps construct. The 1114 bp product obtained was cut with BamHI and EcoRI and cloned into the vector pUC19 (see *Molecular Cloning, A Laboratory Manual*, ibid). The DNA sequence was confirmed as before and the construct was cloned into pET5c vector as a NdeI to EcoRI fragment to generate the plasmid pEt5c HIS 340 scAb.

The jun cassette was assembled by PCR of overlapping oligonucleotides:

Jun1for 5'-atgagaattctcgagcgtatcgctcgtctggaagaaaaagttaaaaccct **(SEQ ID NO: 26)**

Jun 85rev 5'-tagcggtggaagccagttcggagttctgagctttcagggttttaactttt **(SEQ ID NO: 27)**

Jun 71for 5'-tggcttccaccgctaacatgctgcgtgaacaggttgctcagctgaaacag **(SEQ ID NO: 28)**

Jun 146rev 5'-catgcgaattcgtggttcataactttctgtttcagctgagcaacc **(SEQ ID NO: 29)**

The above oligonucleotides were included in a reaction mix at 1pmol each, and the reaction was driven using 10pmol primers Jun 1for-S; 5'-atgagaattctcgagcg **(SEQ ID NO: 30)** and Jun146rev-S; 5'-catgcgaattcgtggttc **(SEQ ID NO: 31)** using high fidelity polymerase and reaction components as previously. The resulting 146bp product was digested with EcoRI, purified and cloned into EcoRI cut pUC19 for sequence analysis using standard procedures (see *Molecular Cloning, A Laboratory Manual* ibid) The Jun cassette was sub-cloned into the pEt5c HIS 340 scAb plasmid as an XhoI-EcoRI fragment by substitution of the existing 320bp XhoI-EcoRI fragment carrying the human constant region domain

Plasmids his-340-jun and FLAG-αPs-fos, were used as templates for PCR using biotinylated primer BioT7; 5'-agatctcgatcccgcaaatta **(SEQ ID NO: 32)** and primer petrev;-5'-aaataggcgtatcacgaggcc **(SEQ ID NO: 33)**. Primers were supplied by GenoSys (Cambridge, UK) and used in the reaction at a concentration of 1pmol. Components and PCR conditions were as previously. The his-340-jun reaction product was 992bp, and the FLAG-αPs-fos reaction product was 1002bp. The products were purified using a spin purification cartridge (Qiagen, Crawley, UK) and diluted to 100ng/µl concentration. Quantitation was by UV absorbance at 260nm. 500ng biotin labelled DNA was reacted with 10µl streptavidin coated magnetic particles (Bangs labs, Fishers, USA). The reaction was conducted in a siliconised microcentrifuge tube in a volume of 500µl PBS 1% (w/v) BSA for 10 minutes at room temperature. Following binding, the particles were collected by magnet (Dynal, Bromborough, UK) and washed three times using PBS 1% BSA.

Please delete the paragraph on page 28, lines 31-37 and replace it with the following paragraph:

In some experiments translation products bound to the particles were detected using antibodies for either the Flag or the his6 **(SEQ ID NO: 34)** epitope engineered into each of the model gene constructs. Antibodies were added to the washed particles diluted in PBS. Incubations were for 60 minutes at 4°C with gentle mixing. A secondary reagent (anti-mouse-HRP conjugate) was added at the recommended dilution in PBS and incubated for a further 30 minutes at 4°C. Particles were washed three times using 200µl PBS before colour development with the chromogenic substrate. Reactions were read at 492nm.

Please delete the paragraph on page 29, lines 1-9 and replace it with the following paragraph:

The presence of the captured target protein gene was confirmed using PCR and DNA sequencing. For detecting the jun model gene, jun specific primers Jun 1for-S and Jun146rev- were used in a PCR assay. The assay was initiated by addition of 10% (v/v) particles directly into the PCR mix. Components and reaction conditions were as previously. The 146bp jun specific product was detected by gel electrophoresis. For detecting the fos model gene, primers Fos1fS and Fos 155rS were used in a PCR assay. Reaction conditions and detection of the 155bp fos specific product were as above. For detecting the negative control protein, primers Seq1scab 5'agatccctactataggta **(SEQ ID NO: 35)** and Seq2scab; 5'-ggtgagctcgatgtatcc **(SEQ ID NO: 36)** were used to detect a 115bp product in the α-Ps scAb protein gene.

Please delete the paragraphs on page 29, line 37 to page 30, line 15 and replace them with the following paragraphs:

In a separate PCR, a linker fragment of form $(Gly_4Ser)_3$ **(SEQ ID NO: 37)** (Huston J. S. et al 1988, *PNAS*, 85: 5879-5883) was amplified from a cloned template pSW1-ScFvD1. 3 (McCafferty et al, 1990, *Nature* **348**: 522-554) using primers sets detailed previously (Marks, J. D in *Antibody Engineering*, ed Borrebaek C.A.K New York O.U.P., 1995). The 93bp linker fragment product was annealed together with an equimolar mixture of the Vh and Vl PCR products. The mixture was further amplified in a "pull through" reaction using flanking primers HuVHBACKsfi and HuFORNot as detailed in Vaughan et al (Vaughan T.J. et al 1996, *Nature Biotech*. **14**: 309-314). All fragments used in the pull-through reaction were purified free of their initial primers prior to inclusion in the reaction. Purification was conducted using the Wizard PCR Preps system from Promega (Promega, Southampton UK).

The assembled contig of form Vh-linker-Vl, was digested with restriction enzymes SfiI and NotI (Boehringer) using standard conditions and purified as above. The purified fragment was annealed with a double stranded synthetic oligonucleotide adapter mix designed to introduce a V8 protease cleavage site juxtaposed with a tract of randomised sequence in frame with the C-terminus of the V1 gene. This V8/unique sequence barcode was produced by annealing a pair of synthetic oligonucleotide pools of form 5'-ggccgcgaggaagaggaa[(atg)/(can)/(agn)/(aan)/(gan)/(ttn)]$_2$gc-3' **(SEQ ID NO: 38)** and 5'-ggccgc[(naa)/(ntc)/(ngt)/(nct)/(nag)/(cat)]$_2$ctccttctcctcgc-3' **(SEQ ID NO: 39)**. This linker has NotI compatible ends (underlined) and therefore facilitates the insertion of the complete

single chain antibody-V8/unique sequence barcode fragment into SfiI-NotI prepared pCANTAB 5 (Pharmacia) phagemid vector.

Please delete the paragraph on page 30, lines 32-39 and replace it with the following paragraph:

For other experiments using *in vitro* transcription and translation (IVTT), the assembled scfv library was subcloned into SfiI NotI prepared pCANTAB5-T7. This vector is the same as the commercially available pCANTAB5 except it was modified to include the T7 promoter sequence (ttaatacgactcactata) **(SEQ ID NO: 40)** inserted at the HindIII site at position 2235. The modification was achieved by ligation of a double-stranded synthetic DNA linker of sequence 5'-agctaatacgactcactata **(SEQ ID NO: 41)** into HindIII cut and de- phosphorylated pCANTAB5. Recombinant clones containing the T7 promoter were selected using a diagnostic PCR.

Please delete the paragraphs on page 33, lines 6-26 and replace them with the following paragraphs:

The C-terminal residues of CHI and a C-terminal FLAG tag sequence (DYKDDDDK) **(SEQ ID NO: 42)** (Knappik A. and Pluckthun A. Biotechniques, 17; 754-761,1990) were added using OL011 and OL012 which included the restriction sites *Eco*lCRI and *Bst*98I in order to produce pC5A8-03. Alternatively, these tags could include the 6HIS tag **(SEQ ID NO: 34)** or MS tags (see example).

The oligonucleotides utilised in the production of pC5A8-0, pC5A8-02 and pC5A8-03 are listed below;

OL001; 5' GGGCAGATCTTTAACTTTAAGAAGGAGATATACATATGAAATACCTATTGCCTACGG 3' (SEQ ID NO: 43)
OL002; 5' GGGTCTGGGTCATAACGATATCGGCCATCGCTGGTTGGGCAGC 3' (SEQ ID NO: 44)
OL003; 5' GGTACCAAACTGGAGATCAAACGGACTGTGGCTGCACCATCT 3' (SEQ ID NO: 45)
OL004; 5' AGATGGTGCAGCCACAGTCCGTTTGATCTCCAGTTTGGTACC 3' (SEQ ID NO: 46)
OL005; 5' GATCGAATTCCTAACACTCTCCGCGGTTGAAGCTCTTTG 3' (SEQ ID NO: 47)
OL006; 5' GATCGAATTCTAACTTTAAGAAGGAGATATACATATG 3' (SEQ ID NO: 48)
OL007; 5' GGACTGAACCAGTTGGACTTCGGCCATCGCTGGTTGGGCAGC 3' (SEQ ID NO: 49)
OL008; 5' ACCCTGGTTAcCGTCTCCTCAGCCTCCACCAAGGGCCCATC 3' (SEQ ID NO: 50)
OL009; 5' GATGGGCCCTTGGTGGAGGCTGAGGAGACGGTAACCAGGGTAC 3' (SEQ ID NO: 51)
OL010; 5' GATCGAGCTCTGCTTTCTTGTCCACCTTGGTGTTGC 3' (SEQ ID NO: 52)

OL011; 5' CCCAAATCTTGCGCTGCAGACTACAAAGACGACGACGACAAATAGCTCGAGC 3' **(SEQ ID NO: 53)**
OL012; 5' TTAAGCTCGAGCTATTTGTCGTCGTCGTCTTTGTAGTCTGCAGCGCAAGATTTGGG 3' **(SEQ ID NO: 54)**

Please delete the paragraph on page 34, lines 13-16 and replace it with the following paragraph:

OL013; 5' GAAGACGTCGCTGTTTAC 3' **(SEQ ID NO: 55)**
OL014; 5' GGTACCAAGCTTGAGATC 3' **(SEQ ID NO: 56)**
OL015; 5' CTACTGCGCGCGTGAAAAAG 3' **(SEQ ID NO: 57)**
OL016; 5' GGGTCAGGGGACCCTGG 3' **(SEQ ID NO: 58)**

Please delete the paragraph on page 34, lines 37-43 and replace it with the following paragraph:

Positive strand; 5'
GAAGACGTCGCTGTTTACTACTGCCAGCAG*NNSNNSNNSNNSNNSNNSNNS*ACCTTCG
GTGGTGGTACCAAGCTTGG 3' **(SEQ ID NO: 59)**

Negative stand: 5'
CCAAGCTTGGTACCACCACCGAAGGT*SNNSNNSNNSNNSNNSNNSNN*CTGCTGGCAGT
AGTAAACAGCGACGTCTTC 3' **(SEQ ID NO: 60)**

Please delete the paragraph on page 35, lines 6-12 and replace it with the following paragraph:

Positive strand; 5'
CTACTGCGCGCGT*NNSNNSNNSNNSNNSNNSNNSNNSNNSNNS*TTCGCTTACTGGGGT
CAGGGGACCCCT **(SEQ ID NO: 61)**

Negative stand: 5'
AGGGGTCCCCTGACCCCAGTAAGCGAA*SNNSNNSNNSNNSNNSNNSNNSNNSNNSNNA*
CGCGCGCAGTAG 3' **(SEQ ID NO: 62)**

Please delete the paragraphs on page 35, line 31 to page 36, line 6 and replace them with the following paragraphs:

As an example a tag of 8 residues can be created using the oligonucleotide 5' NAC NCC NGG NTG TKC VAG GNV CNT 3' **(SEQ ID NO: 2)**. The length of this Tag is increased to 11 residues if a second tag of 8 residues is also included due to the incorporation of the site for protease Factor Xa, which is shown in italics. This allows the tags to be identified as tag 1 or tag 2 following their removal and analysis by mass spectroscopy.

Single tag.

Forward Oligo; 5' GCG CTG CAG *GAY GGN CGN* NAC NCC NGG NTG TKC VAG GNV CNT TAG CTC GAG CTA 3' **(SEQ ID NO: 63)**

Reverse Oligo; 5' TAG CTC GAG CTA ANG BNC CTB GMA CAN CCN GGN GTN CCG CCC GTC CTG CAG CGC 3' **(SEQ ID NO: 64)**

Double tag.

Forward Oligo; 5' GCG CTG CAG *GAY GGN CGN* NAC NCC NGG NTG TKC VAG GNV CNT *GAY GGN CGN* NAC NCC NGG NTG TKC VAG GNV CNT TAG CTC GAG CTA 3' **(SEQ ID NO: 65)**

Reverse Oligo; 5'TAG CTC GAG CTA ANG BNC CTB GMA CAN CCN GGN GTN CCG CCC GTC ANG BNC CTB GMA CAN CCN GGN GTN CCG CCC GTC CTG CAG CGC 3' **(SEQ ID NO: 66)**

Please delete the paragraph on page 37, lines 12-16 and replace it with the following paragraph:

Positive stand; 5' GG GCA GAT CTT TAA CTT TAA GAA GGA GAT ATA CAT ATG AAA TAC CTA TTG CCT ACG G 3' **(SEQ ID NO: 43)**

Negative strand; 5' TAG CTC GAG CTA ANG BNC CTB GMA CAN CCN GGN GTN CCG CCC GTC ANG BNC CTB GMA CAN CCN GGN GTN CCG CCC GTC CTG CAG CGC 3' **(SEQ ID NO: 66)**

Please delete the paragraph on page 37, lines 12-16 and replace it with the following paragraph:

The Pharmacia Recombinant Phage Antibody System (Pharmacia) was used to produce a library of mouse single chain Fvs (ScFv). First-strand cDNA was generated from the mRNA using M-MuLV reverse transcriptase and random hexamer primers. Antibody heavy and light chain genes were then amplified using specific heavy and light chain primers complementary to conserved sequences flanking the antibody variable domains. The 340 and 325 base pair products generated for heavy and light chain DNA respectively were separately purified following agarose gel electrophoresis. These were then assembled into a single ScFv construct using a DNA linker-primer mix to give the VH region joined by a (Gly4Ser)3 **(SEQ ID NO: 37)** peptide to the VL region. The assembled ScFv were amplified with primers designed to insert Sfi 1 and Not 1 sites at the 5' and 3' ends respectively, giving an 800 bp product. This fragment was purified, sequentially digested with SfiI and NotI, and repurified. The fragment was then ligated into SfiI and NotI cut pCANTAB 5 phagemid vector. PCANTAB 5 contains the gene encoding the Phage Gene 3 protein (g3p) and the ScFv is inserted adjacent to the g3 signal sequence such that it will be expressed as a g3p fusion protein. Competent *E. coli* TG1 cells were transformed with the pCantab 5/ScFv phagemid then subsequently infected with the M13KO7 helper phage. The resulting recombinant phage contained DNA encoding the ScFv genes and displayed one or more copies of recombinant antibody as fusion proteins at their tips.